Numerical Scheme for the Generalised Serre-Green-Naghdi Model

Jordan Pitt[a,*], Christopher Zoppou[a], Stephen Roberts[a]

[a]*Australian National University, Canberra*

Abstract

The recently derived generalised Serre-Green-Naghdi equations are investigated and the parameter values classified according to the dispersive properties of the linearised equations with example numerical solutions shown. A numerical scheme for solving the generalised Serre-Green-Naghdi equations is then described. This numerical scheme is an extension to the numerical scheme for the classical Serre-Green-Naghdi equations written in conservation law form. A typical second-order implementation of this numerical scheme is then described and validated using known analytic solutions; the travelling solitary wave and the dam-break problem. The numerical method is further validated for general members of this family of equations using forced solutions. The described numerical method is shown to be conservative, robust and second-order accurate supporting the parameter classification based on the linear theory.

1. Introduction

The generalised Serre-Green-Naghdi Equations (gSGNE) were recently derived by Clamond and Dutykh [1]. They are a family of equations that generalise the classical Serre-Green-Naghdi Equations (SGNE) first derived by Serre [2]. The equations describe the behaviour of water waves in shallow water where the typical water depth h_0 is much smaller than the wavelength λ so that the shallowness parameter $\sigma = h_0/\lambda \ll 1$. The gSGNE are of particular interest to the wave modelling community as their dispersion relationship well approximates the dispersion relationship of the linear wave theory [3]. The choice of two free parameters β_1 and β_2 allows for the dispersion relationship of the gSGNE to be accurate up to $\mathcal{O}\left(\sigma^2\right)$ terms, $\mathcal{O}\left(\sigma^4\right)$ terms and even $\mathcal{O}\left(\sigma^6\right)$ terms [1, 4]. Therefore, these equations provide a way to study the behaviour of water waves as higher powers of σ are retained in the wave model.

Since this family of equations contains dispersive and non-dispersive members, this family of equations also provides a way to study the inclusion of dispersion in non-linear shallow water equations. This has been the objective underpinning much of the current research in the area of dispersive wave modelling. To accomplish this goal multiple investigations have been performed with various numerical methods such as finite element methods [4], discontinuous Galerkin methods [5, 6, 7], hybrid finite volume methods [8, 9] and splitting methods [10, 11, 12] on a number of different dispersive wave equations. This has left the underlying question of the effect of dispersion on non-linear shallow water wave equations answered in a piecemeal way with comparisons of numerical solutions with different numerical methods applied to different equations. To provide a more focused answer to this question we propose one numerical method for an entire family of equations to comprehensively investigate the effect of dispersion on non-linear shallow water equations.

In previous work we have developed and validated numerical methods for a conservative reformulation of the SGNE [13, 9, 14, 15] given by $\beta_1 = 2/3$ and $\beta_2 = 0$ in the gSGNE. These numerical methods have all been based on the following approach, firstly solve an auxiliary elliptic equation for the non-conservative quantity. Secondly, evolve the equations in conservation law form using the finite volume method which requires the non-conserved quantity. The benefit of the approach was its ability to maintain the fundamental conservation properties of the SGNE [15] and the robustness of the method in the presence of steep gradients [16]. The approach has been shown to produce the theoretical accuracy of the underlying methods [14, 15] with the desired conservation and linear dispersion properties [15]. It was demonstrated by Clamond and Dutykh [1] that the gSGNE can also be

---

*jpapitt@gmail.com

written in conservation law form and thus the technique described above for the SGNE can be extended to the gSGNE with arbitrary β_1 and β_2.

This paper begins by introducing the gSGNE and highlighting their important properties. Of particular note, is that the free parameters β_1 and β_2 provide control of the width and relative location of the dispersive wave train. This allows for the family of equations to be classified by their respective dispersive wave train limits, and provides the basis for previously unobserved behaviours in dispersive wave trains. The second-order extension of the numerical scheme of Zoppou et al. [14] is then described.

The numerical method is then validated against analytic solutions of the SGNE and the Shallow Water Wave Equations (SWWE) which are solutions of the gSGNE for the appropriate β values, demonstrating its convergence rate and conservation properties. Forced solutions are also used to validate that all terms in the gSGNE are being approximated to the correct order of accuracy with arbitrary β_1 and β_2 values. Forced solutions are necessary to validate the numerical method for other members of this family of equations, because no analytic solutions are currently known. Together these validations demonstrate the capability of the numerical scheme to produce robust and accurate numerical solutions.

2. Generalised Serre-Green-Naghdi Equations

The gSGNE were derived by Clamond and Dutykh [1] using a Lagrangian field theory approach. These equations generalise the SGNE that describe a depth averaged approximation to the Euler equations, where $h(x,t)$ is the water depth, $u(x,t)$ is the depth averaged horizontal velocity and g is the acceleration due to gravity. This generalisation is achieved by introducing two free parameters β_1 and β_2, each pair of β values corresponding to one member of this family of equations. In particular, the family of equations captured by gSGNE contains the SWWE, the SGNE, a family of regularised SWWE studied by Clamond and Dutykh [1] and a family of improved dispersion SGNE equations studied by Clamond et al. [4].

The gSGNE describe the conservation of mass (h), momentum (uh) and energy $(\mathcal{E})$ for free surface water waves subject to gravitational forces like so

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0 \tag{1a}$$

$$\frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{1}{2}gh^2 + \frac{1}{2}h^2\Gamma\right) = 0 \tag{1b}$$

$$\frac{\partial (\mathcal{E})}{\partial t} + \frac{\partial}{\partial x}\left[hu\left(\frac{1}{2}u^2 + \frac{1}{4}\beta_1 h^2 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x} + gh\left(1 + \frac{1}{4}\beta_2 \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right) + \frac{1}{3}h\Gamma\right) + \frac{1}{2}\beta_2 gh^3 \frac{\partial h}{\partial x}\frac{\partial u}{\partial x}\right] = 0 \tag{1c}$$

where

$$\Gamma = \beta_1 h\left[\frac{\partial u}{\partial x}\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x \partial t} - u\frac{\partial^2 u}{\partial x^2}\right] - \beta_2 g\left[h\frac{\partial^2 h}{\partial x^2} + \frac{1}{2}\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right] \tag{1d}$$

$$\mathcal{E} = \frac{1}{2}hu^2 + \frac{1}{4}\beta_1 h^3 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x} + \frac{1}{2}gh^2\left(1 + \frac{1}{2}\beta_2 \frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right). \tag{1e}$$

When $\beta_1 = \beta_2 = 0$ these equations reduce to the SWWE and the SGNE are recovered when $\beta_1 = 2/3$ and $\beta_2 = 0$.

Equation (1) contains three partial differential equations with only two unknowns, h and u. Since there are only two unknowns it is sufficient to solve only two of the three equations, typically either the conservation of mass (1a) and the conservation of momentum equations (1b) or the conservation of mass (1a) and the conservation of energy equations (1c). If h and u are sufficiently smooth, then the solution of any pair of these equations will produce identical results. However for non-smooth solutions of h and u these pairs of equations will produce different results [17]. Since the mass (1a) and momentum (1b) equations are commonly solved this results in dissipation of energy around discontinuities in the solution. Conversely, when the mass (1a) and energy equations (1c) are solved this results in dissipation of momentum and an incorrect propagation speed for discontinuities in the solution.

When solutions are sufficiently smooth all equations hold simultaneously and the total amounts of all these quantities remain constant in time if the system is closed. This can be seen by integrating (1) over the domain,

and observing that the temporal derivative of the spatial integrals of mass (h), momentum (uh) and energy $(\mathcal{E})$ is zero when there is no flux across the domain boundaries. This conservation property of the gSGNE will be used to validate the numerical method and its solution.

2.1. Linearised gSGNE

The linear dispersion properties of water wave equations have been of particular interest [4, 11, 18], as the scope of wave modelling expands to include dispersive effects. Indeed, the gSGNE are especially relevant because they provide better dispersion properties when compared to the linear theory for water waves [3] than the SWWE where $\beta_1 = \beta_2 = 0$.

The linearised gSGNE (1) are obtained by considering small perturbations, $h_1(x,t)$ in depth and $u_1(x,t)$ in velocity with a background mean water depth h_0 and a mean velocity u_0. By using these assumptions and neglecting all terms significantly smaller than the perturbations we arrive at the linearised gSGNE

$$\frac{\partial h_1}{\partial t} + \frac{\partial}{\partial x}\left(u_0 h_1 + h_0 u_1\right) = 0 \tag{2a}$$

$$\frac{\partial u_1}{\partial t} + \frac{\partial}{\partial x}\left(u_0 u_1 + g h_1 - \frac{\beta_1}{2} h_0^2 \left[\frac{\partial^2 u_1}{\partial x \partial t} + u_0 \frac{\partial^2 u_1}{\partial x^2}\right] - \frac{\beta_2}{2} g h_0^2 \frac{\partial^2 h_1}{\partial x^2}\right) = 0. \tag{2b}$$

The presence of the β_2 term and the choice of β_1 values distinguishes the linearised gSGNE from the linearised SGNE [14].

2.1.1. Dispersion Relation

The dispersion relationship of the linearised gSGNE is obtained by seeking non-trivial travelling wave solutions of the form $\exp\left(i(kx - \omega t)\right)$ of (2) for $h_1(x,t)$ and $u_1(x,t)$. By substituting in the assumed solution form, we obtain the following relationship between the frequency ω in time and the wavenumber k in space

$$\omega_{\text{gSGNE}}^{\pm} = u_0 k \pm k\sqrt{gh_0}\sqrt{\frac{\beta_2 h_0^2 k^2 + 2}{\beta_1 h_0^2 k^2 + 2}}. \tag{3}$$

The dispersion relation has a positive and negative branch corresponding to the direction of these waves. This dispersion relation (3) is equivalent to the dispersion relation derived by Clamond and Dutykh [1] for the gSGNE who assumed that $u_0 = 0$.

The dispersion relationship for the linear water wave theory [3] is

$$u_0 k \pm \sqrt{gk \tanh\left(kh_0\right)}. \tag{4}$$

The power series approximations to the linear dispersion relationship of the gSGNE (3) and the linear water wave theory (4) are

$$\omega_{\text{gSGNE}}^{\pm} = \left(u_0 \pm \sqrt{gh_0}\right)k \pm \frac{\left(\beta_2 - \beta_1\right)}{4}\sqrt{gh_0}\,h_0^2 k^3 \pm \frac{\left(3\beta_1^2 - 2\beta_2\beta_1 - \beta_2^2\right)}{32}\sqrt{gh_0}\,h_0^4 k^5 + \mathcal{O}\left(k^7\right) \tag{5a}$$

$$\omega_{\text{linear}}^{\pm} = \left(u_0 \pm \sqrt{gh_0}\right)k \pm \frac{-1}{6}\sqrt{gh_0}\,h_0^2 k^3 \qquad \pm \frac{19}{360}\sqrt{gh_0}\,h_0^4 k^5 \qquad + \mathcal{O}\left(k^7\right). \tag{5b}$$

Comparing (5a) and (5b), allows the accuracy of the gSGNE dispersion relationship to be determined. In particular, the gSGNE are k^2 accurate for all β values, accurate to the k^4 term when $\beta_1 = \beta_2 + 2/3$ and accurate up to the k^6 term when $\beta_1 = \beta_2 + 2/3$ and $\beta_2 = 2/15$ [4].

From the dispersion relation (3), the phase speed v_p and the group speed v_g are

$$v_p^{\pm} = \frac{\omega_{\text{gSGNE}}^{\pm}}{k} = u_0 \pm \sqrt{gh_0}\sqrt{\frac{\beta_2 h_0^2 k^2 + 2}{\beta_1 h_0^2 k^2 + 2}}, \tag{6a}$$

$$v_g^{\pm} = \frac{\partial \omega_{\text{gSGNE}}^{\pm}}{\partial k} = u_0 \pm \sqrt{gh_0}\sqrt{\frac{\beta_2 h_0^2 k^2 + 2}{\beta_1 h_0^2 k^2 + 2}}\left[1 + \frac{\beta_2 - \beta_1}{\left(\frac{1}{2}\beta_2 h_0^2 k^2 + 1\right)\left(\left(\beta_1 - \frac{1}{3}\right)h_0^2 k^2 + 1\right)}\right]. \tag{6b}$$

The numerical methods for the SGNE described by Le Métayer et al. [8] and Zoppou et al. [14] required that the phase speeds were bounded. To extend these methods to the gSGNE we must demonstrate that the phase speeds of the gSGNE are also bounded.

To demonstrate that the phase speeds are bounded, observe that when $\beta_1 \geq 0$, $\beta_2 \geq 0$ and $h_0 k \geq 0$ then

$$f(h_0 k) = \frac{\beta_2 (h_0 k)^2 + 2}{\beta_1 (h_0 k)^2 + 2},$$

is a monotone function over $h_0 k$. This can be seen by reformulating and taking the derivative with respect to $h_0 k$, to obtain that

$$\frac{\partial (f(h_0 k))}{\partial (h_0 k)} = [\beta_2 - \beta_1] \frac{4 (h_0 k)}{\left(\beta_1 (h_0 k)^2 + 2 \right)^2}.$$

The derivative is greater than zero and thus $f(h_0 k)$ is monotone non-decreasing if $\beta_1 \leq \beta_2$. Whilst the derivative is greater than zero and thus $f(h_0 k)$ is monotone non-increasing if $\beta_1 \geq \beta_2$. Since $v_p^+ = u_0 + \sqrt{g h_0 f(h_0 k)}$ and $v_p^- = u_0 - \sqrt{g h_0 f(h_0 k)}$, given the above properties of $f(h_0 k)$ under the initial assumptions v_p^+ is monotone non-decreasing and v_p^- is monotone non-increasing when $\beta_1 \leq \beta_2$. Conversely, when $\beta_1 \geq \beta_2$ then v_p^+ is monotone non-increasing and v_p^- is monotone non-decreasing.

In addition to the monotonicity of $v_p^\pm$, when $k \to 0$ then $v_p^\pm \to u_0 \pm \sqrt{g h_0}$ and as $k \to \infty$ then $v_p^\pm \to u_0 \pm \sqrt{g h_0} \sqrt{\beta_2/\beta_1}$. Therefore, $v_p^\pm$ is monotonic and bounded at the limits of the domain, and thus bounded for all β values provided that $\beta_1 = 0$ when $\beta_2 = 0$, otherwise in the $k \to \infty$ limit the phase speed is no longer bounded. Consequently, the methods of Le Métayer et al. [8] and Zoppou et al. [14] for the SGNE can be extended to the gSGNE because the phase speeds are bounded.

In addition, to the phase speed bounds we also have the following chain of inequalities when $\beta_1 \geq \beta_2$

$$u_0 - \sqrt{g h_0} \leq v_p^- \leq u_0 - \sqrt{g h_0} \sqrt{\frac{\beta_2}{\beta_1}} \leq u_0 \leq u_0 + \sqrt{g h_0} \sqrt{\frac{\beta_2}{\beta_1}} \leq v_p^+ \leq u_0 + \sqrt{g h_0}. \tag{7}$$

We designate this region of β values, as 'Region 1'. In this region the minimum and maximum phase speed values are $u_0 - \sqrt{g h_0}$ and $u_0 + \sqrt{g h_0}$ respectively. Region 1 is then characterised by trailing dispersive waves when $\beta_2 < \beta_1$ with an upper boundary where waves are not dispersive when $\beta_2 = \beta_1$. Region 1 includes the SWWE and the SGNE and is consistent with the behaviour of the dispersive waves given by the linear theory for water waves [3].

When $\beta_2 > \beta_1$ the inequality chain becomes

$$u_0 - \sqrt{g h_0} \sqrt{\frac{\beta_2}{\beta_1}} \leq v_p^- \leq u_0 - \sqrt{g h_0} \leq u_0 \leq u_0 + \sqrt{g h_0} \leq v_p^+ \leq u_0 + \sqrt{g h_0} \sqrt{\frac{\beta_2}{\beta_1}} \tag{8}$$

This will be denoted as 'Region 2' and it is characterised by advancing dispersive waves. Since we are only interested in water waves with trailing dispersive waves, this paper will be restricted to studying waves in Region 1, although the numerical scheme is valid for all regions.

The regions and location of important members of the gSGNE in terms of β values is summarised in Figure 1. All equations in the gSGNE family are accurate up to $\mathcal{O}(k^2)$ in the dispersion relationship (5). Furthermore, there is a family of equations that obtain $\mathcal{O}(k^4)$ accuracy in the dispersion relationship. Only one particular member of this family, the improved SGNE, obtains a dispersion relationship which is $\mathcal{O}(k^6)$ accurate. Additionally, there are illustrative numerical solutions that demonstrate the effect of the β values and hence the different dispersion relationships on solutions to the dam-break problem described in Section 4.2. The dam-break problem consists of initially still water with a left water depth of $2m$ and a right water depth of $1m$ with a discontinuous jump between the two. The numerical solutions have been well resolved using 25600 cells and the simulated water depth h at $t = 35s$ is plotted. The same numerical solver has been used for each with the only change being the β parameters.

The numerical solutions for the SWWE compare well to the analytic solutions for the SWWE in Section 4.2, and the numerical solutions of the SGNE compare well to the study of the dam-break problems performed

by Pitt et al. [16] for the SGNE. By comparing the numerical solution to its location in the β plot we can see that the behaviour of solutions is described well by the chain of inequalities (7) when $\beta_1 \geq \beta_2$ and (8) when $\beta_1 < \beta_2$. In particular, for the SWWE and the regularised SWWE $\beta_1 = \beta_2$ which results in a phase speed which is independent of k and thus there is no dispersion. Consequently, we observe no wave train between the shock and the rarefaction fan for these non-dispersive wave models. For the dispersive wave models such as the classical SGNE and the improved SGNE with $\mathcal{O}(k^4)$ and $\mathcal{O}(k^6)$ accuracy in the dispersion relationship respectively, we observe a dispersive wave train between the shock and the rarefaction fan. In Region 2 of the β plot we observe advancing dispersive wave trains behind the rarefaction fan and ahead of the shock due to the change of the chain of inequalities when $\beta_1 < \beta_2$ in (8).

2.2. Alternative Conservative Form of the gSGNE

Clamond and Dutykh [1] provided a reformulation of (1b) for the gSGNE, analogous to the reformulation of the SGNE [14, 8, 19]. The purpose of this reformulation is to remove the mixed spatial-temporal derivative in the flux term of the momentum equation (1b), which is difficult to treat numerically. This reformulation is obtained by introducing a new conserved quantity

$$G = hu - \frac{\beta_1}{2}\frac{\partial}{\partial x}\left(h^3\frac{\partial u}{\partial x}\right)$$

and thus (1b) can be written in conservation law form for G as

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x}\left(uG + \frac{gh^2}{2} - \beta_1 h^3 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x} - \frac{1}{2}\beta_2 gh^2\left[h\frac{\partial^2 h}{\partial x^2} + \frac{1}{2}\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right]\right) = 0$$

When $\beta_1 = 2/3$ then G is the same conserved quantity introduced for the SGNE [14, 8, 19].

This reformulation, provides the gSGNE in conservation law form for both h and the new conserved quantity G. Thus the conservative gSGNE are

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0 \tag{9a}$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x}\left(uG + \frac{gh^2}{2} - \beta_1 h^3 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x} - \frac{1}{2}\beta_2 gh^2\left[h\frac{\partial^2 h}{\partial x^2} + \frac{1}{2}\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right]\right) = 0. \tag{9b}$$

with

$$G = uh - \frac{\beta_1}{2}\frac{\partial}{\partial x}\left(h^3\frac{\partial u}{\partial x}\right). \tag{9c}$$

Now that the gSGNE are written in conservation law form and have a bound on the phase speeds, they can be solved numerically using the hybrid finite volume technique described by Zoppou et al. [14] for the SGNE.

3. Numerical Scheme

The proposed numerical scheme for the gSGNE extends those previously published for the SGNE [14, 8] to allow any β values and the associated additional terms. We begin the description of the numerical scheme by giving a brief outline of the finite volume method and the discretisation therein. We then provide an overview of the numerical technique and then describe the simplest second-order implementation of this scheme in detail. With previous studies [14, 15] showing the sufficiency and necessity of second-order methods for the SGNE.

3.1. Finite Volume Method and Discretisation

The core of these hybrid finite volume methods for the gSGNE is the finite volume method used to solve the equations in conservation law form. The finite volume method solves equations of the form

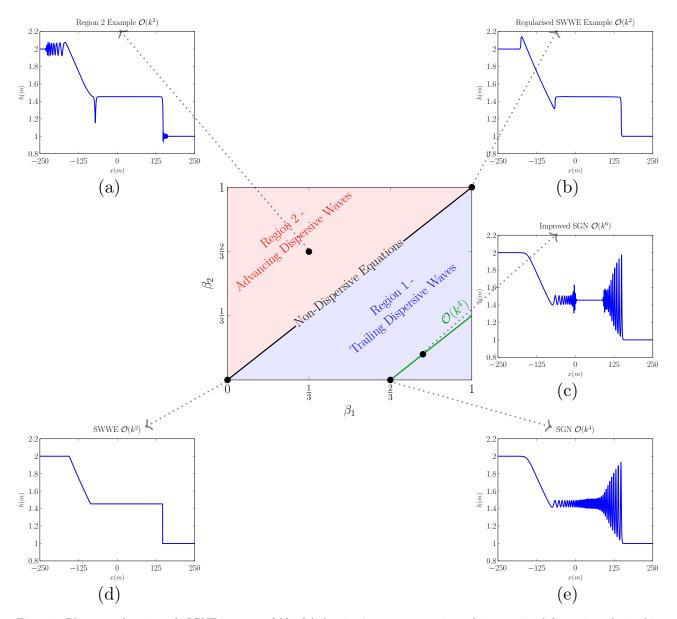$$\frac{\partial q}{\partial t} + \frac{\partial f(q)}{\partial x} = 0 \tag{10}$$

Figure 1: Phase speed regions of gSGNE in terms of (β_1, β_2) showing important equations, their associated dispersion relationship accuracy and an example numerical solution for the dam-break problem solved using $\beta_1 = 1/3$ and $\beta_2 = 2/3$ in (a), $\beta_1 = \beta_2 = 1$ in (b), $\beta_1 = 2/3 + 2/15$ and $\beta_2 = 2/15$ in (c), $\beta_1 = \beta_2 = 0$ in (d) and $\beta_1 = 2/3$ and $\beta_2 = 0$ in (e).

where q is a generic conserved quantity. This is the form of the gSGNE after the reformulation (9). In the finite volume method (10) is integrated over cells of fixed width Δx in space and over time steps of fixed length Δt. The midpoint of the jth cell is given by $x_j = x_0 + j\Delta x$ while the cell edges of the jth cell are given by $x_{j-1/2} = x_0 + (j - \frac{1}{2})\Delta x$ and $x_{j+1/2} = x_0 + (j + \frac{1}{2})\Delta x$. Likewise the nth time step is given by $t^n = t^0 + n\Delta t$.

Integrating (10) over both space and time results in

$$\bar{q}_j^{n+1} = \bar{q}_j^n + \frac{\Delta t}{\Delta x} \left[F_{j+1/2}^n - F_{j-1/2}^n \right] \tag{11}$$

where

$$\bar{q}_j^n = \frac{1}{\Delta x} \int_{x_j - \Delta x/2}^{x_j - \Delta x/2} q(x, t^n) \, dx$$

is the average of q over the jth cell at time t^n and

$$F_{j\pm1/2}^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(q(x_{j\pm1/2}, t)) dt$$

is the average flux of q across the cell edge from time t^n to t^{n+1}. Therefore, if $F_{j\pm1/2}^n$ can be approximated with the appropriate order of accuracy, then we have an explicit method for updating the cell average of the conserved quantities through time for equations in conservation law form (10).

3.2. Overview

We will now describe the numerical scheme, which uses the above finite volume formulation to solve (9a) and (9b). When describing this broad overview it is also useful to consider collections of the point-wise values q or cell averaged values $\bar{q}$ described above over the whole domain at a particular time. Thus for q, we define $\boldsymbol{q}^n$ to be the vector of q_j^n values and $\bar{\boldsymbol{q}}^n$ to be the vector of $\bar{q}_j^n$ values for all cells in the domain at time t^n.

The numerical scheme for the gSGNE based on the finite volume method (11) then proceeds as follows. Begin at the nth time step where the vectors of cell averages of the conserved quantities $\bar{\boldsymbol{h}}^n$ and $\bar{\boldsymbol{G}}^n$ are either known or provided as initial conditions.

1. Solve (9c) using $\bar{\boldsymbol{h}}^n$ and $\bar{\boldsymbol{G}}^n$ to obtain an approximation to $\bar{\boldsymbol{u}}^n$, which can be expressed as

$$\mathcal{A}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n\right) \to \bar{\boldsymbol{u}}^n.$$

2. Solve (9a) and (9b), using the finite volume method (11) with an approximate Riemann solver such as the one described by Kurganov et al. [20], to obtain $\bar{h}_j^{n+1}$ and $\bar{G}_j^{n+1}$ at the next time step, like so

$$\mathcal{F}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n, \bar{\boldsymbol{u}}^n\right) \to \bar{\boldsymbol{h}}^{n+1}, \bar{\boldsymbol{G}}^{n+1}.$$

3. Steps 1 and 2 are repeated and combined using a convex combination to obtain a Strong Stability Preserving (SSP) Runge Kutta time stepping method [21] to approximate $\bar{\boldsymbol{h}}^{n+1}$ and $\bar{\boldsymbol{G}}^{n+1}$ with the correct accuracy in time.

This numerical scheme produces the numerical methods of Le Métayer et al. [8], Zoppou et al. [14] and Pitt [15] when $\beta_1 = 2/3$ and $\beta_2 = 0$. Additionally, when $\beta_1 = \beta_2 = 0$ the gSGNE reduces to the SWWE, and this numerical scheme reduces to a traditional finite volume method [22].

3.3. Typical Second-Order Implementation

The description of this numerical method will be broken up into the steps described in Section 3.2 for simplicity and also to highlight the interchangeability of the different parts of the scheme.

3.3.1. Step 1 - Solution of the Elliptic Equation

To solve (9c) given $\bar{\boldsymbol{h}}^n$ and $\bar{\boldsymbol{G}}^n$ to obtain $\bar{\boldsymbol{u}}^n$ we use the observation that the cell average and the cell nodal value are equal up to second-order accuracy, so that $\bar{\boldsymbol{q}}^n = \boldsymbol{q}^n + \mathcal{O}\left(\Delta x^2\right)$ for all the quantities of interest. Assuming that u is sufficiently smooth, then a second-order central finite difference approximation can be used to accurately solve (9c).

The second-order central finite difference approximation to (9c) can be written as

$$\boldsymbol{u}^n = \boldsymbol{A}^{-1}\boldsymbol{G}^n \tag{12}$$

where $\boldsymbol{A}$ is a tri-diagonal matrix with the sub-diagonal, diagonal and super-diagonal elements are given by

$$A_{j,j-1} = -\frac{\beta_1}{2}\left[\frac{\left(h_j^n\right)^3}{\Delta x^2} - \frac{3\left(h_j^n\right)^2}{2\Delta x}\frac{h_{j+1}^n - h_{j-1}^n}{2\Delta x}\right] \tag{13a}$$

$$A_{j,j} = h_j^n + \beta_1\frac{\left(h_j^n\right)^3}{\Delta x^2} \tag{13b}$$

$$A_{j,j+1} = -\frac{\beta_1}{2}\left[\frac{\left(h_j^n\right)^3}{\Delta x^2} + \frac{3\left(h_j^n\right)^2}{2\Delta x}\frac{h_{j+1}^n - h_{j-1}^n}{2\Delta x}\right] \tag{13c}$$

respectively.

The finite difference approximation (12) can be solved with any matrix solver. We use the explicit Thomas Algorithm [23]. This is an efficient method provided $\boldsymbol{A}$ is diagonally dominant which is the case as long as $h_j^n > 0$, which is true for all the problems of interest in this paper.

Thus as desired we have

$$\mathcal{A}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n\right) = \boldsymbol{A}^{-1}\boldsymbol{G}^n = \boldsymbol{u}^n \to \bar{\boldsymbol{u}}^n. \tag{14}$$

3.3.2. Step 2 - Finite Volume Method - Reconstruction

The finite volume method of Kurganov et al. [20] relies on approximations of all the terms in the spatial derivative of the flux at the cell edges $x_{j\pm1/2}$. Thus the following quantities require second-order approximations at the cell edges: u, h, G, $\partial u/\partial x$, $\partial h/\partial x$, $\partial^2 h/\partial x^2$. In this paper we are interested in validating the numerical method by reproducing solutions to smooth analytic and forced solutions for all members of the gSGNE. We are also interested in validating the numerical method for the discontinuous dam-break solution for only the SWWE where $\beta_1 = \beta_2 = 0$ and thus only h, u and $G = uh$ need to be approximated. For this reason our gSGNE solver uses approximations to u, h, G that allow for discontinuities and thus require limiting. Whereas the approximations to the derivatives $\partial u/\partial x$, $\partial h/\partial x$, $\partial^2 h/\partial x^2$ will assume that these quantities are smooth and thus do not require limiting.

Figure 1 demonstrates the adequacy of this reconstruction approach even in situations where the derivatives are present in the equations but are not smooth. A main reason for this is that the elliptic equation (9c) helps control the smoothness of u. However, this does not explain why we get reasonable results when assuming the derivatives of h are smooth when they are not smooth in the initial conditions. For the validations performed in this paper, the assumption on the smoothness of the derivatives $\partial u/\partial x$, $\partial h/\partial x$, $\partial^2 h/\partial x^2$ is adequate.

The reconstructions for u, h and G will be summarised for a general quantity q. For the jth cell we have the reconstruction at the left boundary $q_{j-1/2}^+ = \bar{q}_j - s_j\Delta x/2$ and the reconstruction at the right boundary $q_{j+1/2}^- = \bar{q}_j + s_j\Delta x/2$. The slope of the reconstruction s_j is calculated like so

$$s_j = \text{minmod}\left(\theta\frac{\bar{q}_j - \bar{q}_{j-1}}{\Delta x}, \frac{\bar{q}_{j+1} - \bar{q}_{j-1}}{2\Delta x}, \theta\frac{\bar{q}_{j+1} - \bar{q}_j}{\Delta x}\right). \tag{15}$$

The minmod function is defined as follows

$$\text{minmod}\left(a, b, c\right) = \begin{cases} \min\left(a, b, c\right) & \text{when} \quad a > 0, b > 0, c > 0 \\ \max\left(a, b, c\right) & \text{when} \quad a < 0, b < 0, c < 0 \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

This gives a reconstruction at each cell edge, so cell j gives rise to $q_{j-1/2}^+$ and $q_{j+1/2}^-$ while cell $j+1$ gives rise to $q_{j+1/2}^+$ and $q_{j+3/2}^-$, all of which are required in the flux approximation.

For $\partial u/\partial x$, $\partial h/\partial x$ and $\partial^2 h/\partial x^2$ we assume that the quantities are smooth, and so using the appropriate order finite difference approximation at the cell edges is sufficient. We demonstrate these approximations for $x_{j+1/2}$ and a generic quantity q

$$\left[\frac{\partial q}{\partial x}\right]_{j+1/2} = \frac{q_{j+1} - q_j}{\Delta x},$$

$$\left[\frac{\partial^2 q}{\partial x^2}\right]_{j+1/2} = \frac{q_{j+2} - q_{j+1} - q_j + q_{j-1}}{2\Delta x^2}.$$

3.3.3. Step 2 - Finite Volume Method - Flux Approximation

To solve (9a) and (9b) which are both in conservation law form (10) we use (11). Since we begin with $\bar{\boldsymbol{h}}^n$ and $\bar{\boldsymbol{G}}^n$, we only require an approximation to $F^n_{j\pm 1/2}$ to obtain $\bar{\boldsymbol{h}}^{n+1}$ and $\bar{\boldsymbol{G}}^{n+1}$.

In this numerical scheme the approximate Riemann solver described by Kurganov et al. [20] is used to calculate $F^n_{j\pm 1/2}$. The major advantage of this scheme is that it only requires bounds on the phase speeds. Only the calculation of the flux term $F^n_{j+1/2}$ is demonstrated as the process to calculate the flux term $F^n_{j-1/2}$ is identical but with different cells. For q the flux is approximated by

$$F^n_{j+\frac{1}{2}} = \frac{a^+_{j+\frac{1}{2}} f\left(q^-_{j+\frac{1}{2}}\right) - a^-_{j+\frac{1}{2}} f\left(q^+_{j+\frac{1}{2}}\right)}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} + \frac{a^+_{j+\frac{1}{2}} a^-_{j+\frac{1}{2}}}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} \left[q^+_{j+\frac{1}{2}} - q^-_{j+\frac{1}{2}}\right] \tag{17}$$

where $a^+_{j+\frac{1}{2}}$ and $a^-_{j+\frac{1}{2}}$ are given by the phase speed bounds and all quantities on the right hand side are computed at time t^n. Applying the phase speed bounds (7) and (8) we obtain

$$a^-_{j+\frac{1}{2}} = \min\left\{0 \,,\, u^-_{j+1/2} - \max\left(1, \sqrt{\frac{\beta_2}{\beta_1}}\right)\sqrt{gh^-_{j+1/2}} \,,\, u^+_{j+1/2} - \max\left(1, \sqrt{\frac{\beta_2}{\beta_1}}\right)\sqrt{gh^+_{j+1/2}}\right\}, \tag{18}$$

$$a^+_{j+\frac{1}{2}} = \max\left\{0 \,,\, u^-_{j+1/2} + \max\left(1, \sqrt{\frac{\beta_2}{\beta_1}}\right)\sqrt{gh^-_{j+1/2}} \,,\, u^+_{j+1/2} + \max\left(1, \sqrt{\frac{\beta_2}{\beta_1}}\right)\sqrt{gh^+_{j+1/2}}\right\} \tag{19}$$

where the $\max\left(1, \sqrt{\beta_2/\beta_1}\right)$ accounts for the different phase speed bounds (7) and (8), which depend on the choice of β values.

The flux functions $f(q^-_{j+\frac{1}{2}})$ and $f(q^+_{j+\frac{1}{2}})$ are evaluated using the reconstructed values of the jth and $(j+1)^{th}$ cell respectively. From the continuity equation (9a) we have

$$f\left(h^\pm_{j+\frac{1}{2}}\right) = u^\pm_{j+1/2} h^\pm_{j+1/2}.$$

For the evolution of G equation (9b) we have

$$f\left(G^\pm_{j+\frac{1}{2}}\right) = u^\pm_{j+1/2} G^\pm_{j+1/2} + \frac{g}{2}\left(h^\pm_{j+1/2}\right)^2 - \beta_1 \left(h^\pm_{j+1/2}\right)^3 \left(\left[\frac{\partial u}{\partial x}\right]_{j+1/2}\right)^2$$

$$- \frac{1}{2}\beta_2 g \left(h^\pm_{j+1/2}\right)^2 \left[h^\pm_{j+1/2}\left[\frac{\partial^2 h}{\partial x^2}\right]_{j+1/2} + \frac{1}{2}\left(\left[\frac{\partial h}{\partial x}\right]_{j+1/2}\right)^2\right] \tag{20}$$

where the derivatives are assumed to be smooth across the cell edge, and thus do not require different superscripts.

Since the reconstructions were given above for all these quantities, we can approximate $F^n_{j\pm 1/2}$ using (17) and thus employ (11) to obtain $\bar{\boldsymbol{h}}^{n+1}$ and $\bar{\boldsymbol{G}}^{n+1}$ resulting in

$$\mathcal{F}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n, \bar{\boldsymbol{u}}^n\right) = \bar{q}^n_j + \frac{\Delta t}{\Delta x}\left[F^n_{j+1/2} - F^n_{j-1/2}\right] \rightarrow \bar{\boldsymbol{h}}^{n+1}, \bar{\boldsymbol{G}}^{n+1} \tag{21}$$

as desired.

3.3.4. Step 3 - Runge-Kutta Time Stepping

Combining Steps 1 and Steps 2, provides a spatially second-order scheme with first-order time stepping. To arrive at a fully second-order method, we repeat Steps 1 and Steps 2 according to the second-order SSP Runge Kutta method [21]. Doing this we obtain the following scheme making use of the above implementation of $\mathcal{A}$, (14) and $\mathcal{F}$, (21)

$$\bar{\boldsymbol{h}}^{(1)}, \bar{\boldsymbol{G}}^{(1)} = \mathcal{F}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n, \mathcal{A}\left(\bar{\boldsymbol{h}}^n, \bar{\boldsymbol{G}}^n\right)\right)$$

$$\bar{\boldsymbol{h}}^{(2)}, \bar{\boldsymbol{G}}^{(2)} = \mathcal{F}\left(\bar{\boldsymbol{h}}^{(1)}, \bar{\boldsymbol{G}}^{(1)}, \mathcal{A}\left(\bar{\boldsymbol{h}}^{(1)}, \bar{\boldsymbol{G}}^{(1)}\right)\right)$$

$$\bar{\boldsymbol{h}}^{n+1}, \bar{\boldsymbol{G}}^{n+1} = \frac{1}{2}\left(\bar{\boldsymbol{h}}^n + \bar{\boldsymbol{h}}^{(2)}\right), \frac{1}{2}\left(\bar{\boldsymbol{G}}^n + \bar{\boldsymbol{G}}^{(2)}\right).$$

obtaining a conditionally stable fully second-order method for solving (9).

3.4. Boundary Conditions

For the purposes of the validation below and for simplicity, we have applied Dirichlet boundary conditions using ghost cells over which the values of h, G and u are known. Thus in addition to the m cells inside the boundary we have an additional l cells either side of it so that we have the left ghost cells $-l$, $-l+1$, ..., -1 and the right ghost cells m, $m+1$, ..., $m+l-1$. Since the numerical method has a maximum stencil that extends 2 cells beyond the target cell, then l must be greater than 1.

This boundary condition is applied to Step 1 by extending the matrix equation (12) to vectors containing the ghost cells $\hat{\boldsymbol{u}}^n$ and $\hat{\boldsymbol{G}}^n$ like so

$$\hat{\boldsymbol{u}}^n = \begin{bmatrix} u_{-l}^n & \dots & u_{-1}^n & u_0^n & \dots & u_{m-1}^n & u_m^n & \dots & u_{m+l-1}^n \end{bmatrix}^T$$

$$\hat{\boldsymbol{G}}^n = \begin{bmatrix} G_{-l}^n & \dots & G_{-1}^n & G_0^n & \dots & G_{m-1}^n & G_m^n & \dots & G_{m+l-1}^n \end{bmatrix}^T.$$

Likewise the associated extended matrix $\hat{\boldsymbol{A}}$ in (12) is the same as $\boldsymbol{A}$ when $0 \le j \le m-1$ (13) with the additional ghost cell values given by

$$\hat{A}_{j,j-1} = 0 \qquad \hat{A}_{j,j} = 1 \qquad \hat{A}_{j,j+1} = 0$$

when $j < 0$ or $j > m-1$.

For Step 2 there is no need to reconstruct the quantities inside the ghost cells because h, G and u are given.

3.5. Courant-Frederichs-Lewy Condition

To ensure the stability of the finite volume method (11) the Courant-Friedrichs-Lewy (CFL) condition [24] is used. The CFL condition is necessary for stability and ensures that time steps are small enough so that information is only transferred between neighbouring cells. For the gSGNE the CFL condition is

$$\Delta t \le \frac{Cr}{\max_j \left\{a_{j+1/2}^{\pm}\right\}} \Delta x \tag{22}$$

where $a_{j+1/2}^{\pm}$ are the phase speed bounds used in the flux approximation (19) and $0 \le Cr \le 1$ is the Courant number.

4. Validation

The numerical method described above is validated using analytic solutions for particular β values that correspond to the SGNE and the SWWE as well as a forced solution. Together these tests demonstrate the ability of the method to reproduce analytic solutions to important members of the gSGNE family, as well as assessing the accuracy of the numerical method's approximation of all terms in the gSGNE.

4.1. Convergence and Conservation Measures

To validate the numerical solutions we make use of measures of convergence and conservation. The measure of convergence will be the relative deviation of the numerical solution from the equivalent analytic or forced solution using the L_2 norm. While the conservation properties of the numerical method will be measured by numerically approximating the energy in the initial conditions and the numerical solution and comparing them using a measure called C_1.

For a quantity q with a vector of its analytic or forced solution at the cell midpoints $\boldsymbol{q}$ and the numerical solution at the cell midpoints $\boldsymbol{q}^*$, the discrete non-dimensional L_2 norm is

$$L_2\left(\boldsymbol{q}, \boldsymbol{q}^*\right) = \left(\frac{\sum_j \left(q_j - q_j^*\right)^2}{\sum_j q_j^2} \right)^{\frac{1}{2}} \tag{23}$$

where the time-step superscripts were suppressed for simplicity.

For a quantity q with a vector of its values at the nth time step $\boldsymbol{q}^n$, the total amount of the quantity is approximated by $C(\boldsymbol{q}^n)$. The method for this is the same as the method described by Zoppou et al. [14], which has a higher order of accuracy than the numerical method used to solve the gSGNE. Using the numerical approximation to the total amounts, the conservation error is obtained using

$$C_1\left(\boldsymbol{q}^0, \boldsymbol{q}^n\right) = \begin{cases} \dfrac{\left|C(\boldsymbol{q}^0) - C(\boldsymbol{q}^n)\right|}{\left|C(\boldsymbol{q}^0)\right|} & , \quad \left|C(\boldsymbol{q}^0)\right| > 0 \\[2ex] \left|C(\boldsymbol{q}^0) - C(\boldsymbol{q}^n)\right| & , \quad \left|C(\boldsymbol{q}^0)\right| = 0 \end{cases} . \tag{24}$$

4.2. Analytic Solutions

The analytic solutions used to validate the numerical method, are the solitary travelling wave solution of the SGNE and the dam-break solution of the SWWE. The solitary travelling wave solution is a smooth travelling wave solution, where the non-linear and dispersive terms are in balance in the gSGNE. Whereas the dam-break solution of the SWWE demonstrates the robustness of the method in the presence of steep gradients.

4.2.1. SGNE - Solitary Travelling Wave Solution

When $\beta_1 = 2/3$ and $\beta_2 = 0$ the gSGNE are equivalent to the SGNE which admit the following travelling wave solution [25]

$$h(x,t) = a_0 + a_1 \mathrm{sech}^2\left(\kappa(x - ct)\right), \tag{25a}$$

$$u(x,t) = c\left(1 - \frac{a_0}{h(x,t)}\right), \tag{25b}$$

where

$$\kappa = \frac{\sqrt{3a_1}}{2a_0\sqrt{a_0 + a_1}} \tag{25c}$$

and

$$c = \sqrt{g\left(a_0 + a_1\right)} \tag{25d}$$

is the speed of the wave.

This travelling wave solution is maintained due to a balance between the dispersive terms and the non-linear terms in the momentum equation (1b). This balance results in a solitary wave that is advected with a constant speed without a changing shape. Validating the numerical solutions for the gSGNE solver using this solution tests the balance between these terms in (1b), and allows us to verify the method's conservation of energy as the solution is smooth. To enable a comparison between the numerical method and the SGNE solver of Zoppou et al. [14] and Pitt [15], the chosen solitary travelling wave parameters were $a_0 = 1m$ and $a_1 = 0.7m$ and the acceleration due to gravity at the Earth's surface, $g = 9.81m^2/s$ was used.
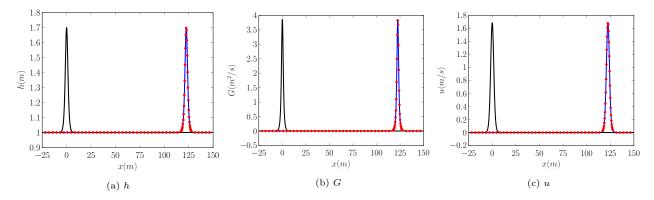
(a) h (b) G (c) u

Figure 2: Plot of comparing initial ($\blacksquare$), analytic solution ($\blacksquare$), and numerical solution with $\Delta x \approx 0.06m$ ($\bullet$) to the solitary travelling wave analytic solution of the gSGNE when $\beta_1 = 2/3$ and $\beta_2 = 0$ at $t = 30s$.
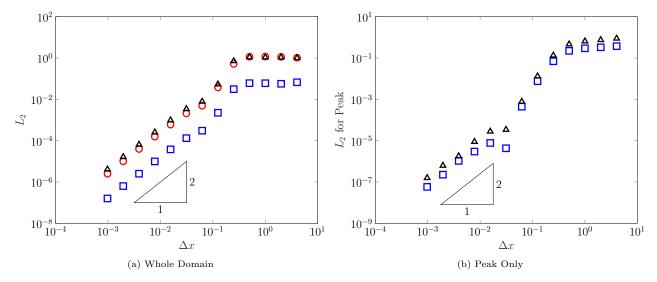

(a) Whole Domain   (b) Peak Only

Figure 3: Convergence plots for h ($\square$), G ($\bigcirc$) and u ($\triangle$) for the numerical solutions of the gSGNE when $\beta_1 = 2/3$ and $\beta_2 = 0$ as Δx varies.

The numerical solution was solved over the domain $[-200m, 200m]$ from $t = 0s$ until $t = 30s$. The spatial resolution was varied like so $\Delta x = 400/(100 \times 2^l)$, where l was increased from 0 to 12. To satisfy the CFL condition (22) the time step $\Delta t = \Delta x/(2\sqrt{g(a_0 + a_1)})$ was chosen. The limiting parameter $\theta = 1.2$, was chosen to match previous numerical experiments [14, 15].

Example numerical solutions for h, u and G with $\Delta x = 400/(100 \times 2^6) \approx 0.06m$ are plotted in Figure 2. These examples demonstrate that the numerical solutions can reproduce the analytic solutions well.

The convergence properties of the method as Δx varies are given in Figure 3 where Figure 3a shows the L_2 norm calculated over the whole domain and Figure 3b shows the L_2 norm at the peak only. The L_2 norm over the whole domain demonstrates that all quantities are approximated with second-order accuracy, and thus the method is second-order accurate. The L_2 norm at the peak was measured by taking the relative difference between the numerical solutions value at the peak of the wave which are given by $\max(\boldsymbol{h}^n)$, $\max(\boldsymbol{u}^n)$ to the analytic values which are $a_0 + a_1$ and $a_0 c/(a_0 + a_1)$ respectively. The L_2 norm at the peak demonstrates that even with the slope limiting and the diffusion introduced by the flux approximation [16], the numerical method introduces little diffusion and is second-order accurate.

The conservation error C_1 (24) of the numerical solutions as Δx varies is given in Figure 4. This conservation measure demonstrates that the finite volume method conserves h and G up to round-off error, which accumulates as Δx increases due to the additional number of calculations. Since uh and $\mathcal{E}$ are not conserved quantities in our finite volume method which solves (9a) and (9b), they are not conserved up to round-off error. However,
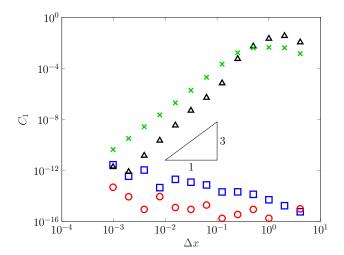
Figure 4: Conservation plot for h ($\square$) , G ($\bigcirc$), uh ($\triangle$) and $\mathcal{E}$ ($\times$) for the numerical solutions of the gSGNE when $\beta_1 = 2/3$ and $\beta_2 = 0$ as Δx varies.
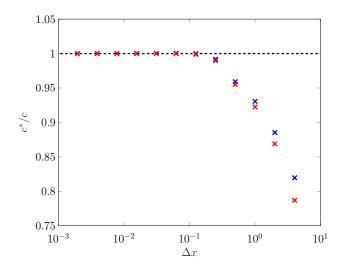

Figure 5: Plot of c_+^*/c ($\times$), c_-^*/c ($\times$) and the analytic value (– –) as Δx varies.

the conservation properties of uh and $\mathcal{E}$ are good, with both exhibiting conservation errors that have better than second-order convergence in Δx.

By locating the cell on which $\boldsymbol{h}^n$ achieves its maximum, we can provide an upper and lower bound for the speed of the numerical solution to the travelling wave problem where the lower bound is $c_-^* = (x_{\text{peak cell}} - 0.5\Delta x)/t$ and the upper bound is $c_+^* = (x_{\text{peak cell}} + 0.5\Delta x)/t$. These upper and lower bounds are compared to the analytic value c given by (25d) in Figure 5. This figure demonstrates the diffusion of the numerical scheme when Δx is large because both the upper and lower bounds on the wave speed are below the analytic value. However, this diffusion becomes negligible when Δx is small, and for the lowest Δx value we observe that $c_-^* < c < c_+^*$, indicating that the peak is travelling at the correct speed up to cell width accuracy.

These results agree well with the numerical solutions of Pitt [15], who compared various numerical methods for the SGNE. This demonstrates that the numerical method solving the gSGNE when $\beta_1 = 2/3$ and $\beta_2 = 0$ accurately reproduces the analytic solution of the SGNE.

4.2.2. SWWE - Dam-break Solution

When $\beta_1 = \beta_2 = 0$ the gSGNE equations reduce to the SWWE. Consequently, (9b) and (1b) become identical with $G = uh$. The SWWE possess an analytic solution to the dam-break problem given by the initial

conditions

$$h(x,0) = \begin{cases} h_0 & x < 0 \\ h_1 & x \geq 0 \end{cases} \tag{26a}$$

$$u(x,0) = 0 \tag{26b}$$

$$G(x,0) = 0. \tag{26c}$$

The solution to the dam-break problem using the conservation of mass (1a) and momentum (1b) equations is given by

$$h(x,t) = \begin{cases} h_0 & , & x \leq -t\sqrt{gh_0} \\ \frac{4}{9g}\left(\sqrt{gh_0} - \frac{x}{2t}\right)^2 & , & -t\sqrt{gh_0} < x \leq t\left(u_2 - \sqrt{gh_2}\right) \\ h_2 & , & t\left(u_2 - \sqrt{gh_2}\right) < x \leq tS \\ h_1 & , & tS \leq x \end{cases} \tag{27a}$$

$$u(x,t) = \begin{cases} 0 & , & x \leq -t\sqrt{gh_0} \\ \frac{2}{3}\left(\sqrt{gh_0} + \frac{x}{t}\right) & , & -t\sqrt{gh_0} < x \leq t\left(u_2 - \sqrt{gh_2}\right) \\ u_2 & , & t\left(u_2 - \sqrt{gh_2}\right) < x \leq tS \\ 0 & , & tS_2 \leq x \end{cases}. \tag{27b}$$

The constant state values h_2 and u_2 and the shock speed S can be calculated for any initial conditions by solving

$$h_2 = \frac{h_0}{2}\left(\sqrt{1 + 8\left(\frac{2h_2}{h_2 - h_0}\left(\frac{\sqrt{gh_1} - \sqrt{gh_2}}{\sqrt{gh_0}}\right)\right)^2} - 1\right) \tag{28a}$$

$$u_2 = 2\left(\sqrt{gh_1} - \sqrt{gh_2}\right), \tag{28b}$$

$$S = \frac{2h_2}{h_2 - h_1}\left(\sqrt{gh_0} - \sqrt{gh_2}\right). \tag{28c}$$

The initial conditions as well as the analytic solution are discontinuous. Due to the discontinuities the solutions to the initial conditions are not unique, as solving the momentum equation (1b) or the energy equation (1c) results in solutions with similar structures but different shock speeds. The solution presented above is solution of the mass (1a) and momentum (1b) equations, as these equations are the basis of the numerical method.

A number of numerical experiments were run for the dam-break problem with $h_0 = 2m$ and $h_1 = 1m$. The domain of the solution was $[-250m, 250m]$ with a final time of $t = 35s$. The spatial resolution was varied like so $\Delta x = 500/(100 \times 2^l)$ where l was increased from 0 to 12. To satisfy the CFL condition (22) the time step length $\Delta t = \Delta x/\left(2\sqrt{gh_0}\right)$ was used. The limiting parameter $\theta = 1.0$ and the acceleration due to gravity $g = 9.81m^2/s$ were used.

Example numerical solutions with the spatial resolution $\Delta x = 500/(100 \times 2^5) \approx 0.15m$ and the analytic solutions for h, G and u at the final time are plotted in Figure 6. These figures demonstrate that the method is robust in the presence of steep gradients and accurately reproduces the analytic solution.

The presence of discontinuities in the analytic solution makes accurately assessing convergence of the numerical solutions difficult. To circumvent these issues, the convergence measure has been restricted to comparing the numerical and analytic solutions for the constant region between the rarefaction fan and the shock. This modified convergence measure as Δx is varied is plotted in Figure 7a. This figure demonstrates that the scheme retains it's second-order accuracy away from discontinuities, as desired.

Since the analytic solution contains discontinuities all three conservation laws for h, G and $\mathcal{E}$ are not all satisfied simultaneously [17]. Since we solved equations for h and G, these quantities are conserved in the analytic solutions however, $\mathcal{E}$ is no longer conserved and energy is lost as the shock propagates. Therefore, we introduce a new measure $\mathcal{E}^*$ which measures the conservation error of the energy by comparing the total energy in the numerical solution with the total energy in the analytic solution at the final time. The conservation error for mass, momentum and energy as calculated normally and the new corrected energy are all compared
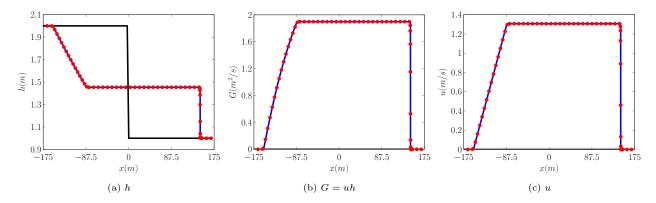
(a) h (b) $G = uh$ (c) u

Figure 6: Comparison of initial (━), analytic solution (━), and numerical solution with $\Delta x \approx 0.15m$ (●) to the dam-break problem for the gSGNE when $\beta_1 = \beta_2 = 0$ at $t = 35s$.

(a) L_2 for constant state for h (□), G (○) and u (△)   (b) C_1 for h (□), $G = uh$ (○), $\mathcal{E}$ (✖) and $\mathcal{E}^*$ (◇)
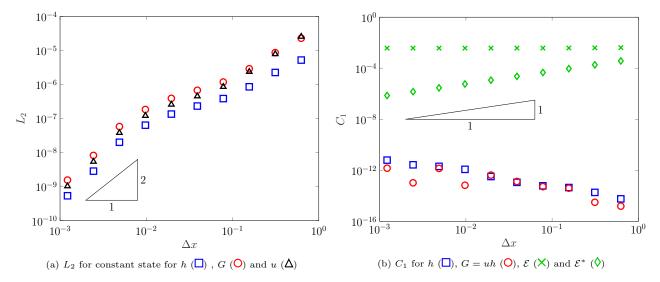
Figure 7: Convergence and conservation plots for the numerical solutions of the gSGNE when $\beta_1 = \beta_2 = 0$ as Δx varies.

in Figure 7b. This figure demonstrates that due to the use of the finite volume method even in the presence of discontinuities the conserved quantities h and G are conserved up to round-off error, leading to their conservation error increasing as Δx decreases. Energy is not conserved by the analytic solution, and this can be seen as the conservation error of $\mathcal{E}$ does not improve as Δx decreases. However, when accounting for the lack of energy conservation by using $\mathcal{E}^*$ we are able to recover first-order accuracy. Given that this solution contains a discontinuity these are good conservation results [26].

To further justify the ability of the method to resolve the discontinuous analytic solution, lower x_{lower} and upper x_{upper} bounds for the location of the shock in the numerical solution were found. These lower and upper bounds for the location are then used to bound the shock speed in the numerical solution. For x_{lower} this was accomplished by finding the first cell with $\bar{h}_j \leq h_1 + \frac{9}{10}(h_2 - h_1)$ and for x_{upper} this was achieved by finding the first cell with $\bar{h}_j \leq h_1 + \frac{1}{10}(h_2 - h_1)$. Consequently, the lower bound for the shock speed $S_-^* = x_{\text{lower}}/t$ and the upper bound for the shock speed $S_+^* = x_{\text{upper}}/t$ were calculated and compared to the analytic value S given by (28c) in Figure 8. This figure demonstrates that as Δx decreases the numerical solutions better resolve the shock. Furthermore it demonstrates that x_{lower} and x_{upper} provide a bound for the true location of the shock.

These results demonstrate that the analytic solution of the SWWE has been accurately reproduced by the numerical method which solves the gSGNE with $\beta_1 = \beta_2 = 0$.
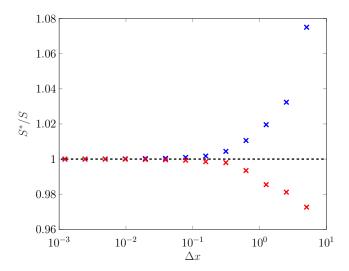
Figure 8: Plot of S_+^*/S ($\times$), S_-^*/S ($\times$) and analytic value (— —) as Δx varies.

4.3. Forced Solutions

There are currently no known analytic solutions to the gSGNE for other β values. Hence, to demonstrate the validity and versatility of the method to solve the gSGNE for other β values, forced solutions are necessary. This is vital for the gSGNE in particular because for the β values tested above, β_2 is always zero, and thus the known analytic solutions do not assess the numerical methods accuracy for the β_2 term.

To generate a forced solution the forced gSGNE are considered

$$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} = \frac{\partial h^*}{\partial t} + \frac{\partial (u^*h^*)}{\partial x} \tag{29a}$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x}\left(uG + \frac{gh^2}{2} - \beta_1 h^3 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x} - \frac{1}{2}\beta_2 gh^2\left[h\frac{\partial^2 h}{\partial x^2} + \frac{1}{2}\frac{\partial h}{\partial x}\frac{\partial h}{\partial x}\right]\right) =$$
$$\frac{\partial G^*}{\partial t} + \frac{\partial}{\partial x}\left(u^*G^* + \frac{g\left(h^*\right)^2}{2} - \beta_1\left(h^*\right)^3\frac{\partial u^*}{\partial x}\frac{\partial u^*}{\partial x} - \frac{1}{2}\beta_2 g\left(h^*\right)^2\left[h^*\frac{\partial^2 h^*}{\partial x^2} + \frac{1}{2}\frac{\partial h^*}{\partial x}\frac{\partial h^*}{\partial x}\right]\right). \tag{29b}$$

The forced gSGNE admit the solutions h^*, u^* and G^* assuming G^* satisfies (9c). Since these equations are satisfied for any chosen functions for h^*, u^* and G^* and any β values, forced solutions can be used to verify the method for a larger class of problems than permitted by currently known analytic solutions. Since the left hand-side of these modified equations are approximated by the numerical method, combining the numerical method with the analytic expressions for the right hand-side, produces a method that approximates the forced gSGNE (29) with the same convergence properties as the underlying numerical method for the gSGNE.

The following forced solution

$$h^*(x,t) = a_0 + a_1 \exp\left(\frac{(x - a_2 t)^2}{2a_3}\right) \tag{30a}$$

$$u^*(x,t) = a_4 \exp\left(\frac{(x - a_2 t)^2}{2a_3}\right) \tag{30b}$$

where G^* is given by (9c), were used. These forced solutions describe Gaussian bumps in h and u that travel at a constant speed a_2. This forced solution was chosen because it is smooth and the terms in (29) are not constant over the whole domain. Smoothness is necessary for the current description of the forced solutions, since it requires the derivatives to be defined in the classical strong sense. This ensures that the forced solutions include all the terms in the gSGNE when assessing the numerical method.
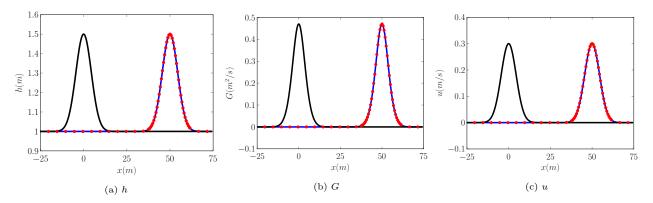
(a) h (b) G (c) u

Figure 9: Plots of the initial conditions (━), analytic solution (━), and numerical solution with $\Delta x \approx 0.06m$ (●) for the forced gSGNE with $\beta_1 = 2/3 + 2/15$ and $\beta_2 = 2/15$.
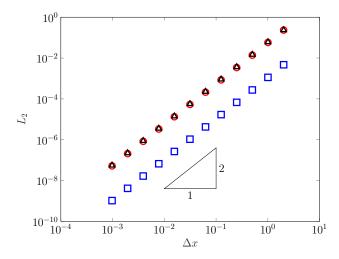

Figure 10: Convergence plot of h (□) , G (○), u (△) for the solutions of the forced gSGNE with $\beta_1 = 2/3 + 2/15$ and $\beta_2 = 2/15$ for various Δx values.

The particular parameter values $a_0 = 1$, $a_1 = 0.5$, $a_2 = 5$, $a_3 = 20$ and $a_4 = 0.3$ were chosen in this investigation, while multiple β values were tested we will be focusing on $\beta_1 = 2/15 + 2/3$ and $\beta_2 = 2/15$ below. This choice of β values corresponds to the improved SGNE which have a $\mathcal{O}\left(k^6\right)$ accurate dispersion relationship. There is an example numerical solution for these equations in Figure 1(c).

The numerical solutions were produced over the domain $[-100m, 100m]$ with a final time of $t = 10s$. The spatial resolution was varied like so $\Delta x = 200/(100 \times 2^l)$ with l varied from 0 to 11. To satisfy the CFL condition the time step size $\Delta t = \Delta x / \left(2\left[a_4 + a_2 + \sqrt{g\left(a_0 + a_1\right)}\right]\right)$ was chosen. The acceleration due to gravity $g = 9.81m^2/s$ was used. For the forced solutions the limiting on the reconstruction on h, u and G was removed. Because these forced solutions are smooth, such reconstruction is not necessary.

Figure 9 compares an example numerical solution at the final time for h, G and u with $\Delta x = 200/(100\times2^5) \approx 0.06m$ with the analytic solution. These example solutions demonstrate that the numerical method is able to reproduce the forced solution well, validating the numerical methods approximation to all terms in (9).

Figure 10 demonstrates the convergence of the numerical scheme as Δx decreases. All quantities of interest are converging at the expected second-order. Since the right hand-sides of (29) are evaluated analytically, the observed error is caused by the numerical method alone. Therefore, these results demonstrate that the scheme is second-order for all terms in the gSGNE.

5. Experimental Comparison for Dispersive Shock Waves

The evolution of an initial rectangular depression wave in the free surface was studied experimentally by Hammack and Segur [27]. The experiments were conducted in a wave tank $0.394m$ wide, $31.6m$ long and $0.61m$ high filled by water $h_0 = 0.1m$ deep. The initial wave was generated by a $0.61m$ long piston with its left edge flush with the left edge of the wave tank. The piston is originally in the up position with the water stationary, the piston is then dropped creating the rectangular depression in the surface, which generates a dispersive shock wave. This dispersive shock wave has its rarefaction fan component propagate up the wave tank, while the front of the dispersive shock wave reflects from the left edge of the wave tank and propagates back up the wave tank. The evolution of these waves is recorded at timeseries of the wave tank is recorded by Wave Gauges (WG) located $0m$, $5m$, $10m$, $15m$ and $20m$ away from the right edge of the piston. Two heights of the rectangular wave were investigated $h_0 - h_1 = 0.01m$ and $h_0 - h_1 = 0.03m$ where h_1 is the height of the water above the piston after it drops.

These experiments provide a good benchmark for the capability of the numerical method to accurately model problems with steep gradients in the free surface. However, these experiments are affected by bed friction, viscosity and the inability of the piston and water to move vertically instantaneously and slip free. The Serre equations do not contain viscosity and bed friction was neglected in this thesis. Additionally, discontinuous initial conditions are used to model the negative rectangular wave. Therefore, it is expected that the numerical solutions of the Serre equations will produce more oscillations in the dispersive wave train than are observed experimentally [16].

Hammack and Segur [27] report the results for two different initial negative wave amplitudes $0.01m$ and $0.03m$, resulting in the non-linearity parameters $\epsilon = 0.1$ and $\epsilon = 0.3$ respectively. Since these non-linearity parameters are relatively small there was no breaking of waves throughout the experiment.

These experiments were modelled numerically using the reflected problem, with the left wall of the wave tank as the axis of symmetry. In the numerical experiments the domain is $[-60m, 60m]$ and the experiment is run for $50s$ with $g = 9.81m/s^2$. The spatial resolution used was $\Delta x = 0.01m$ while the CFL condition (22) was satisfied by using $\Delta t = 0.5\Delta x/\sqrt{g\,0.1}$. The limiting parameter $\theta = 1.2$ was used in the reconstruction (**??**) in FEVM$_2$ and FDVM$_2$.

The numerical results of FDVM$_2$ for these experiments were published by Zoppou et al. [14]. These results have been extended in this thesis with the inclusion of the conservation error of h, G and uh in the numerical simulation of both experiments.

5.1. Results for the $0.01m$ Negative Rectangular Wave

Comparisons of the numerical and experimental WG data for the $0.01m$ negative rectangular wave are displayed in Figures **??** and **??** for FEVM$_2$ and FDVM$_2$ respectively. The data is presented in the same dimensionless scales as reported in the original paper [27]. Furthermore, Tables **??** and **??** display the completely numerically calculated error in conservation C^*.

The numerical solutions have reproduced the experimental results well; particularly for the front of the dispersive wave train. While the total amounts of h, G, uh and $\mathcal{H}$ have been conserved by the numerical methods with little error.

The numerical solutions produce larger and consequently faster waves and generate more oscillations in the dispersive wave train than are observed experimentally. These discrepancies can be attributed to the lack of viscosity and the omission of bed friction for the Serre equations in this thesis (**??**). Furthermore, it is highly likely that the experiment produced some smooth approximation to a discontinuous jump in the water depth with the down-stroke of the piston. Such a smoothing of the initial conditions will significantly attenuate the high frequency waves in the generated dispersive wave train [16]. Given these differences the numerical methods replicated the experimental behaviour well.

These numerical solutions compare well to those of Zoppou et al. [14] for FDVM$_2$ with $\Delta x = 0.005m$, $\Delta t = 0.2\Delta t/\sqrt{g0.1}s$ and $\theta = 1$. The differences in the resolution and the value of θ had little impact on the WG results. However, the more diffusive value of the limiting parameter θ smoothed the initial discontinuity which negatively impacted the conservation of $\mathcal{H}$ reported by Zoppou et al. [14].

Both FEVM$_2$ and FDVM$_2$ have produced indistinguishable results at this scale and have demonstrated very good conservation of all the quantities as demonstrated in Tables **??** and **??**. Given the extensive examination of several numerical schemes for steep gradient problems [16], this indicates that these solutions are representative

$h_0 - h_1$	β_1	β_2	C_1 for h	C_1 for G	C_1 for $\mathcal{E}$
$1cm$	0	0	6.238×10^{-14}	9.648×10^{-20}	4.939×10^{-5}
	2/3	0	8.816×10^{-14}	2.156×10^{-17}	5.898×10^{-6}
	4/5	2/15	8.949×10^{-14}	1.056×10^{-17}	1.579×10^{-5}
$3cm$	0	0	5.286×10^{-14}	3.221×10^{-19}	6.577×10^{-4}
	2/3	0	8.715×10^{-14}	2.106×10^{-17}	1.295×10^{-4}
	4/5	2/15	8.403×10^{-14}	1.528×10^{-18}	2.364×10^{-4}

Table 1: Conservation of conserved quantitites for numerical solutions to all negative rectangular wave experiments.

of true solutions of the Serre equations. However, these results do not demonstrate the superiority of one of these methods over the other.

5.2. Results for the $0.03m$ Negative Rectangular Wave

The WG data for the numerical and experimental results for the evolution of the $0.03m$ negative rectangular wave are displayed in Figures ?? and ?? for FEVM$_2$ and FDVM$_2$ respectively. These results are reported using the same dimensionless scales the original paper [27]. The completely numerically calculated conservation error C^* for all of the conserved quantities are given in Tables ?? and ?? for FEVM$_2$ and FDVM$_2$ respectively.

Both methods reproduce the overall behaviour of this experiment very well. Since this rectangular wave is deeper than the $0.01m$ negative rectangular wave experiment, this experiment provides a more rigorous test of the numerical methods. However, the increased depth also strengthens the effects that produced the discrepancy between the experimental results and the numerical solutions of the Serre equations in the $0.01m$ negative rectangular wave experiment. Hence, the greater differences observed between the numerical and experimental results for the amplitude and speed of the generated waves.

Since the negative rectangular wave is larger than in the $0.01m$ negative rectangular wave experiment, the numerical methods have a larger error in conservation for all the quantities, except h which is conserved exactly. For G and uh the conservation errors are around machine precision and can be disregarded, thus only the error in conservation of $\mathcal{H}$ was significantly increased by a larger negative wave. However, the error in conservation of $\mathcal{H}$ was small and therefore all quantities were well conserved by the numerical methods.

These results compare well with those produced by Zoppou et al. [14] using FDVM$_2$ with $\Delta x = 0.005m$, $\Delta t = 0.2\Delta t/\sqrt{g0.1}s$ and $\theta = 1$. The WG data of those numerical solutions and the solutions displayed here are indistinguishable. The error in conservation in $\mathcal{H}$ is very similar as well, although for the numerical solutions displayed here the numerical grid is coarser. The poorer than expected conservation of $\mathcal{H}$ given the finer grid of the numerical solution is caused by the limiting parameter $\theta = 1$ diffusing the initial discontinuous jump [16].

Conserving $\mathcal{H}$ is difficult for numerical methods solving steep gradient problems [16], especially when two steep gradients interact with one another over short time spans. Increasing the resolution of the numerical method improves the conservation of $\mathcal{H}$ as demonstrated for the analytic solutions. However, the results for the WG data for higher resolution solutions will be indistinguishable from the results presented here and so the provided solutions are sufficient.

These experiments have been replicated equally well by the numerical methods. Given the resolution, the error in conservation and the extensive study summarised in Chapter ??; these results demonstrate the accuracy of the numerical methods in the presence of steep gradients in the free surface.

6. Conclusion

The members of the gSGN family of equations were classified according to the width and relative location of their dispersive wave trains using the dispersion relationship of the linearised equations. Example numerical solutions for select members of this family of equations were used to demonstrate that the full nonlinear gSGN equations exhibit the same behaviour.
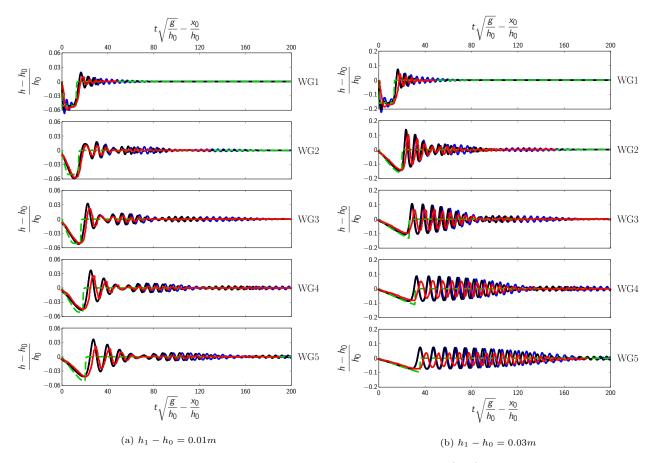
(a) $h_1 - h_0 = 0.01m$

(b) $h_1 - h_0 = 0.03m$

Figure 11: Rectangular negative wave experiments time series at the five Wave Gauges (WG) for simulation with the improved SGN with $\beta_1 = 2/3 + 2/15$ and $\beta_2 = 2/15$ (—), the classical SGN with $\beta_1 = 2/3$ and $\beta_2 = 0$ (—), the SWWE with $\beta_1 = 0$ and $\beta_2 = 0$ (— —) compared to the experimental results (—).

The numerical method used to generate these solutions was then described as a modified version of the numerical scheme for the SGNE outlined by Zoppou et al. [14]. This numerical scheme for the gSGNE was validated by describing a fully second-order implementation of the numerical method. This numerical method was validated against analytic solutions of the SGNE and SWWE and forced solutions. The analytic solutions demonstrate that the gSGNE solver accurately reproduces important members of the family of equations described by the gSGNE whilst conserving the quantities of interest. The forced solutions demonstrate that the method remains second-order for all values of the free parameters, β_1 and β_2. The gSGNE method described above is the first validated numerical method for the gSGNE and its validation supports the provided classification of members of the gSGN using their linear dispersive properties.

References

[1] D. Clamond, D. Dutykh, Non-dispersive conservative regularisation of nonlinear shallow water (and isentropic euler equations), Communications in Nonlinear Science and Numerical Simulation 55 (2018) 237–247.

[2] F. Serre, Contribtion à l'étude des écoulements permanents et variables dans les canaux, La Houille Blanche 6 (1953) 830–872.

[3] G. B. Whitham, Non-linear dispersion of water waves, Journal of Fluid Mechanics 27 (1967) 399–412.

[4] D. Clamond, D. Dutykh, D. Mitsotakis, Conservative modified Serre-Green-Naghdi equations with improved dispersion characteristics, Communications in Nonlinear Science and Numerical Simulation 45 (2017) 245–257.

[5] H. Dong, M. Li, A reconstructed central discontinuous Galerkin-finite element method for the fully nonlinear weakly dispersive Green-Naghdi model, Applied Numerical Mathematics 110 (2016) 110–127.

[6] N. Panda, C. Dawson, Y. Zhang, A. B. Kennedy, J. J. Westerink, A. S. Donahue, Discontinuous Galerkin methods for solving Boussinesq-Green-Naghdi equations in resolving non-linear and dispersive surface water waves, Journal of Computational Physics 273 (2014) 572–588.

[7] M. Li, L. Xu, Y. Cheng, A CDG-FE method for the two-dimensional Green-Naghdi model with the enhanced dispersive property, Journal of Computational Physics 399 (2019) 108953.

[8] O. Le Métayer, S. Gavrilyuk, S. Hank, A numerical scheme for the Green-Naghdi model, Journal of Computational Physics 229 (2010) 2034–2045.

[9] C. Zoppou, S. Roberts, J. Pitt, A solution of the conservation law form of the Serre equations, The Australia and New Zealand Industrial and Applied Mathematics Journal 57 (2016) 385–394.

[10] D. Lannes, F. Marche, A new class of fully nonlinear and weakly dispersive Green-Naghdi models for efficient 2D simulations, Journal of Computational Physics 282 (2015) 238–268.

[11] A. G. Filippini, M. Kazolea, M. Ricchiuto, A flexible genuinely nonlinear approach for nonlinear wave propagation, breaking and run-up, Journal of Computational Physics 310 (2016) 381–417.

[12] J. do Carmo, J. Ferreira, L. Pinto, G. Romanazzi, An improved Serre model: Efficient simulation and comparative evaluation, Applied Mathematical Modelling 56 (2018) 404–423.

[13] C. Zoppou, Numerical Solution of the One-dimensional and Cylindrical Serre Equations for Rapidly Varying Free Surface Flows, Ph.D. thesis, Australian National University, Mathematical Sciences Institute, College of Physical and Mathematical Sciences, Australian National University, Canberra, ACT 2600, Australia, 2014.

[14] C. Zoppou, J. Pitt, S. Roberts, Numerical solution of the fully non-linear weakly dispersive Serre equations for steep gradient flows, Applied Mathematical Modelling 48 (2017) 70–95.

[15] J. Pitt, Simulation of Rapidly Varying and Dry Bed Flow using the Serre equations solved by a Finite Element Volume Method, Ph.D. thesis, Australian National University, Mathematical Sciences Institute, College of Physical and Mathematical Sciences, Australian National University, Canberra, ACT 2600, Australia, 2019.

[16] J. Pitt, C. Zoppou, S. Roberts, Behaviour of the Serre equations in the presence of steep gradients revisited, Wave Motion 76 (2018) 61–77.

[17] Y. Pu, R. Pego, D. Dutykh, D. Clamond, Weakly singular shock profiles for a non-dispersive regularization of shallow-water equations, Communications in Mathematical Sciences 16 (2018) 1361–1378.

[18] J. do Carmo, J. Ferreira, L. Pinto, On the accurate simulation of nearshore and dam break problems involving dispersive breaking waves, Wave Motion 85 (2019) 125 – 143.

[19] M. Li, P. Guyenne, F. Li, L. Xu, High order well-balanced CDG-FE methods for shallow water waves by a Green-Naghdi model, Journal of Computational Physics 257 (2014) 169–192.

[20] A. Kurganov, S. Noelle, G. Petrova, Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton-Jacobi equations, Journal of Scientific Computing, Society for Industrial and Applied Mathematics 23 (2002) 707–740.

[21] S. Gottlieb, C. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, Review, Society for Industrial and Applied Mathematics 43 (2001) 89–112.

[22] S. Roberts, C. Zoppou, Explicit schemes for dam-break simulations, Journal of Hydraulic Engineering 129 (2003) 11–34.

[23] S. Conte, C. De Boor, Elementary numerical analysis: An algorithmic approach, International Series in Pure and Applied Mathematics, 3rd ed., McGraw-Hill Inc., New York, 1980.

[24] R. Courant, K. Friedrichs, H. Lewy, On the partial difference equations of mathematical physics, IBM Journal of Research and Development 11 (1967) 215–234.

[25] G. El, R. Grimshaw, N. Smyth, Unsteady undular bores in fully nonlinear shallow-water theory, Physics of Fluids 18 (2006) 027104.

[26] C. Shu, Efficient algorithms for solving partial differential equations with discontinuous solutions, Notices of the AMS 59 (2012) 615–621.

[27] J. Hammack, H. Segur, The Korteweg-de Vries equation and water waves. Part 3. Oscillatory waves, Journal of Fluid Mechanics 84 (1978) 337–358.